



Sistemas abertos aplicados à gestão da informação científica: o desenvolvimento da UC Digitalis

Ana Miguéis

evamigueis@sib.uc.pt

Ana Luísa Silva

als@uc.pt

Álvaro Trindade

atrindade@sib.uc.pt

Bruno Neves

bneves@sib.uc.pt

Serviço Integrado das Bibliotecas da Universidade de Coimbra

Introdução

O crescente volume de informação, associado à sua desmaterialização, traduz uma complexidade acrescida para as bibliotecas, e para as organizações em geral, na gestão da informação publicada na Web. Também a evolução tecnológica levou à alteração de metodologias e processos de trabalho, o que resultou na necessidade de pôr em prática novas ferramentas de gestão e organização da informação. A implementação de sistemas de gestão de conteúdos (CMS) permite descentralizar todos os processos envolvidos nessa gestão, desde a criação até à publicação, revertendo na otimização dos serviços prestados.

Objetivos

Pretende-se, com este trabalho, dar a conhecer a aplicação de um sistema de gestão de conteúdos no contexto de uma universidade portuguesa. A estratégia global de gestão de informação na Universidade de Coimbra é a de promover a utilização de sistemas interoperáveis entre si e com outros sistemas externos, permitindo assegurar uma gestão de informação adequada, uma racionalização de processos e um aumento da visibilidade da investigação produzida, dentro e fora da instituição. Com este propósito surgiu o projeto UC Digitalis, disponível em <https://digitalis.uc.pt>, constituído por um conjunto de bibliotecas

digitais com recursos digitais, individualizados e complementares, que procura assumir-se como uma plataforma de difusão da cultura e ciência produzidas em espaço lusófono.

A criação desta plataforma surgiu em 2011 e resultou de uma parceria inicial entre dois serviços da Universidade de Coimbra: a Imprensa da Universidade de Coimbra (IUC), que tinha a necessidade de uma plataforma para armazenar e disponibilizar os seus ebooks e revistas; e o Serviço Integrado das Bibliotecas (SIBUC), que dispunha do conhecimento para a desenvolver, fruto de experiências anteriores, nomeadamente dos projectos AlmaMater (<http://almamater.uc.pt>), Estudo Geral (<https://estudogeral.sib.uc.pt>), Classica Digitalia (<http://classica.digitalia.uc.pt>) e portal dos Serviços de Biblioteca e Documentação da Faculdade de Letras da Universidade de Coimbra (<https://alpha.sib.uc.pt>). Mais recentemente conta com a participação do Instituto de Investigação Interdisciplinar (IIIUC) da mesma Universidade.

De início criaram-se as bibliotecas digitais Pombalina e Impactum, sendo a primeira uma biblioteca digital de obras editadas pela Imprensa da Universidade de Coimbra, e a segunda uma biblioteca digital de publicações periódicas, ambas vocacionadas para promover e divulgar a ciência gerada em espaço lusófono. Mais tarde, integrou-se a biblioteca digital de Fundo Antigo da Universidade de Coimbra, Almamater, constituída por um vasto acervo de obras publicadas antes de 1940. Pretendeu-se valorizar os importantes núcleos que integram o rico património bibliográfico e documental das várias bibliotecas, e testemunhar um processo, iniciado em 2010, de disponibilização gratuita e universal de importantes obras do património da Universidade.

Gradualmente o volume de informação disponível nesta plataforma aumentou de modo significativo, disponibilizando, atualmente, perto de 13.000 documentos distribuídos pelas três bibliotecas digitais que a compõem e que correspondem a 25 títulos de revistas em suporte digital com 7.000 artigos; 2.000 monografias publicadas nos últimos cinco anos pela Imprensa da Universidade; e a inclusão de mais de 4.000 obras digitalizadas em projetos que envolveram os fundos antigos das bibliotecas da Universidade de Coimbra e que alimentam a Almamater.

A UC Digitalis apresenta-se, assim, como uma plataforma desenvolvida na UC, mas que se define como um projeto global e prioritário da Universidade de Coimbra para a agregação e difusão de conteúdos digitais.

Metodologia

Durante o planeamento deste projeto foi escolhido o modelo de desenvolvimento de *software* em cascata tendo sido identificadas as tarefas, respetivas precedências e estimado o tempo para a sua conclusão. Foram, depois, assentes as etapas principais do projeto, realizada a sua distribuição pelos intervenientes e definida a data do seu início.

De entre os principais requisitos identificados destacam-se: a acessibilidade da plataforma através da web; a capacidade de pesquisa e agregação de recursos, a existência de uma área pessoal de utilizador, a possibilidade de criar listas de referências bibliográficas e, como requisito não funcional, o uso de sistemas de gestão de conteúdos open source para o desenvolvimento da plataforma. Foi, então, realizado um levantamento das soluções existentes e, das que cumpriam satisfatoriamente os requisitos exigidos, foi feita uma avaliação com base em critérios objetivos. Decidiu-se que a solução mais vantajosa seria usar dois sistemas de gestão de conteúdos, para agilizar o desenvolvimento da plataforma: um com funções de *backoffice*, para gestão de administração dos objetos digitais (Dspace) e outro (Drupal) como *frontend* onde se apresentam os conteúdos e onde o utilizador acede.

Resultados e discussão

A razão da escolha de dois em vez de um CMS resultou da complementaridade que ambos apresentaram quando da sua análise. O Dspace armazena os objetos digitais (texto integral e subprodutos) e apresenta vantagens sobre o Drupal no que respeita à escalabilidade, segurança, performance, implementação procedimentos de preservação digital, indexação por motores de busca de índole académica, etc. O Drupal apresenta vantagens claras no que se refere à capacidade de inclusão e desenvolvimento de novas funcionalidades, devido em parte à sua ampla comunidade.

A interoperabilidade entre os dois CMS foi assegurada através da descrição dos documentos usando o esquema qualificado de metadados Dublin Core, e sua exposição através do Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) com implementação de sets, filtros e modificadores. Foram criados procedimentos que, de forma automática, e a partir do texto integral, geram os subprodutos (capas, *thumbnails*, *previews*, etc). Do lado do Drupal a comunicação com o Dspace utiliza os módulos Feeds e Feeds OAI-PMH Fetcher and Parser que permitem a importação dos dados expostos pelo OAI-PMH e a respetiva agregação por diferentes tipos de conteúdos, de acordo com a tipologia de documento. O módulo Feeds assegura ainda a funcionalidade dos Feeds RSS. Para a indexação e pesquisa de conteúdos por parte dos utilizadores foi usado o Apache Lucene™ Solr. Foi configurado na modalidade *multi-core*, com uma instância global e uma instância para cada uma das bibliotecas digitais que compõem a UC Digitalis, possibilitando pesquisas compartimentadas.

Estão disponíveis funções de pesquisa simples com recurso a facetas e pesquisa avançada, através de taxonomias ou do texto integral. Existem, ainda, funções de ajuda ao preenchimento (*auto complete*) e sugestões de pesquisa alternativas (*did you mean*). O módulo de pesquisa é usado para as funções de navegação (*browsing*) com possibilidade de inclusão de filtros. Disponibiliza, para cada documento e para cada utilizador, diferentes estilos de citação como Chicago, Harvard, APA, NP405, ABNT, etc. Permite ainda a

exportação dos registos em diversos formatos normalizados como o RIS, ENDNOTE, Bibtex e XML.

Aos ebooks e artigos científicos da UCDigitalis foi atribuído um identificador do objeto digital (DOI) da CrossRef que também figura na referência bibliográfica.

A leitura dos conteúdos disponibilizados pode ser feita utilizando um leitor de ebooks, ferramenta que resulta da adaptação do leitor de ebooks open source do Internet Archive. Esta opção permite melhorar a experiência da leitura, aproximando o digital do documento físico e está adaptada para uso em dispositivos tácteis, com funções como o *page turning*.

Relativamente ao aspeto gráfico, a página principal da UC Digitalis permite ligações para a pesquisa, navegação e área pessoal e disponibiliza uma nuvem de etiquetas com as áreas do conhecimento. Existe, também, uma montra ou “carrossel” que de forma dinâmica vai apresentando as últimas novidades e ligações para as bibliotecas Impactum, Pombalina e Almamater que acolhem, respetivamente, revistas e artigos, ebooks, e fundo antigo. A página de cada um dos documentos exhibe os principais metadados, incluindo, no caso dos artigos, o resumo, autoria, citação, capa, *preview*, etc. Permite, ainda, a partilha do documento em redes sociais e o seu descarregamento.

A organização dos conteúdos assenta nas áreas do conhecimento definidas pela Classificação de Domínios Científicos e Tecnológicos da FOS, que se organiza em seis grandes áreas científicas, acessíveis na plataforma através de uma nuvem de etiquetas.

A UC Digitalis está atualmente disponível em 4 idiomas: português (pt), inglês (en), espanhol (es) e mandarim (zh), tendo sido usado o módulo Internationalization (<https://www.drupal.org/project/i18n>) para este fim. Permite não apenas a escolha inicial do idioma, mas também a comutação de idioma em qualquer altura da navegação na plataforma, e a escolha automática do idioma com base nas definições do browser do cliente.

A UC Digitalis enquanto provedora de dados tem implementado um servidor SRW/SRU, que permite a realização de pesquisas federadas por entidades externas (sites, serviços web) sobre os seus conteúdos, e permite as operações *explain* e *searchretrieve* com pesquisas sobre os principais metadados em formato Dublin Core. Oferece, complementarmente, interfaces de comunicação baseados no protocolo OAI-PMH, possibilitando, deste modo, a disponibilização e integração com outros sistemas de informação e agregadores de conteúdos. Suporta a versão 2.0 do OAI-PMH, expondo os dados no formato Dublin Core.

A UC Digitalis recorre à API do Google Analytics (GAPI) para exibir e guardar informação estatística sobre o uso da plataforma. Esta funcionalidade não é pública, estando reservada a administradores. É possível ter acesso a estatísticas sobre o número de

visualizações, visitas, descarregamentos e origem geográfica dos mesmos, num determinado intervalo temporal. Adicionalmente é feita uma recolha periódica do número de citações dos artigos no Google Scholar e guardado o respetivo histórico.

Os utilizadores da UC Digitalis podem autenticar-se através do protocolo open source *Lightweight Directory Access Protocol* (LDAP), estando a autenticação federada com recurso ao protocolo open source Shibboleth em fase de implementação. A UC Digitalis permite, finalmente, a autenticação com base no IP de origem do cliente. Em alternativa a estes métodos o utilizador pode-se autenticar através de login/password, após ter solicitado a criação de conta. De forma a evitar o uso abusivo dos formulários de criação de conta (e outros na plataforma) foi usado o módulo Captcha.

Conclusão

Os resultados atrás referidos evidenciam as vantagens decorrentes do desenvolvimento desta plataforma que resultaram da implementação do projeto, que agrega e difunde conteúdos digitais.

O desenvolvimento de novas funcionalidades, a integração em bases de dados de referência, a indexação de conteúdos na ISI Web of Knowledge (em tarefas de colaboração com o IIIUC), a interoperabilidade com outras plataformas digitais – nomeadamente repositórios institucionais e plataformas editoriais como o Open Journal Systems – são áreas de intervenção que se projetam para um futuro próximo.

A UC Digitalis assume-se, assim, como uma plataforma de difusão de informação e de conhecimento, otimizando os serviços prestados à comunidade, num projeto aberto à colaboração entre a Universidade de Coimbra e o meio envolvente, quer a nível nacional, quer internacional.

Palavras-chave: Sistemas de Gestão de Conteúdos; Drupal; Dspace; Bibliotecas digitais; Projeto UC Digitalis; Universidade de Coimbra

Referências bibliográficas

DRUPAL [Em linha]. [Consult. 23 mar. 2014] Disponível na Internet: <URL:<https://www.drupal.org>>

DSPACE [Em linha]. [Consult. 17 mar. 2014] Disponível na Internet: <URL:<http://www.dspace.org>>.

MIGUÉIS, Ana [et al.] - *UC Digitalis* [Em linha]. Aveiro : [s.n], 2013. [Consult. 23 Março 2014]. Disponível na Internet: <URL: <http://hdl.handle.net/10316/24486>>.

NEVES, Bruno Daniel Pinheiro - *Os Sistemas de Gestão de Conteúdos aplicados à gestão da informação em bibliotecas universitárias* [Em linha]. Coimbra : [s.n], 2011. [Consult. 5 abr. 2014]. Dissertação de mestrado. Disponível na Internet: URL: <<http://hdl.handle.net/10316/14462>>